Introduction

Malaria is among one of the most prominent global diseases, infecting between three-hundred and five-hundred million people yearly, with annual global fatalities averaging over two million (Gagnon 2002). Although the disease is primarily confined to the tropics and subtropics, a resurgence in disease distribution and frequency has been noted since the early 1970s (Patz 1996). Leading researchers in the field of epidemiology, those such as Jonathan Patz and Richard Stone of the American Medical Association, believe that this increase may be the result of climate change, particularly global warming.

As global warming increases, there is an expected rise in the frequency of irregular climate fluctuations, including the phenomenon known as El Niño Southern Oscillation, which causes uncharacteristic changes in atmospheric conditions. Because of its reliance on a tropical genus of mosquitoes as a method of transmission, malaria is particularly sensitive to changes in temperature and precipitation. Any increase in the frequency of El Niño events could bolster epidemics in ranges and capacities history has not yet seen, specifically in less-developed countries where disease abatement is minimal. This fear has brought the linking of El Niño and malaria to the forefront of epidemiology.

In particular, epidemiologists like Menno Jan Bouma of the London School of Health and Tropical Medicine, are conducting research in order to draw correlations between sea-surface temperatures (an early indicator of El Niño) and malaria prevalence.
within specified regions. Such correlations would allow the utilization of sea-surface temperatures as a successful method for forecasting malaria epidemics.

<NOTE: LIT REVIEW NOT INCLUDED>

Methodology

Despite geographical climatopathology’s relative obscurity, it is relied upon heavily within the field of epidemiology as a useful source for understanding the dynamics of disease distribution and seasonal and interannual outbreaks. Many different types of research exist in this field, however the discipline lends itself especially well to statistical applications. The outcome of studies correlating climate and disease are extremely valuable as a diagnostic and analytical tool. Consequently, results can be used to forecast disease distribution through the assessment of climate and climate change. For example, within the past two decades, quite a number of studies have been conducted to investigate the effects of ENSO on malaria epidemics.

As do the climatic effects of El Niño, results of these studies vary drastically depending on geographic location of the study area. The climatic effects of El Niño are not spatially consistent, thus their influence on malaria outbreaks is not consistent across the landscape. In some areas, epidemics respond favorably to flooding as opposed to drought. Often there is no immediate correlation, but rather a lag effect in epidemics following ENSO activity. Yet other times, no correlation is found at all.

Climatic anomalies are not always temporally consistent, either. Large variations among ENSO events may be problematic to the researcher (Lindsay 2000). One ENSO
cycle may vary from another in magnitude and duration. Even the frequency at which El Niño occurs may fluctuate between two and seven years (Kovats 2000). Gagnon stresses the importance of investigating each ENSO event individually. However, this is not necessarily a valuable strategy when attempting to use sea-surface temperature anomalies to detect the probability of an oncoming malaria threat and how severe it may be. A reliable statistical correlation must be found between sea-surface temperatures and malaria prevalence in order to predict the probability of a malaria epidemic occurring.

In order to deem if there was indeed a statistically significant correlation between Eastern Equatorial Pacific sea-surface temperatures and malaria prevalence, a retrospective analysis comparing two datasets was necessary.

Malaria prevalence was assessed using the number of reported malaria cases per year, as acquired by the World Health Organization Epidemiology and Burden of Disease Unit, Global Programme on Evidence for Health Policy Mortality and Morbidity Database. Morbidity alone was chosen, rather than morbidity and mortality, because this was a more accurate assessment of malaria prevalence. Additionally, the number of reported cases of malaria per year did not fluctuate and greatly as mortality rates, and would most likely be a more consistent indicator of epidemic prevalence rather than severity. Malaria mortality in Colombia has decreased greatly in recent decades, whereas malaria morbidity has shown an almost consistent upward trend during the last forty years. Moreover, using morbidity rather than mortality would help to alleviate the effects of medical advancements, which may have confounded results.

Eastern Equatorial Pacific sea-surface temperature datasets were acquired from the National Oceanic and Atmospheric Administration’s National Centers for
Environmental Prediction, Environmental Modeling Center, Climate Modeling Branch. Records were in months, but were computed into 12-month yearly averages (January through December) in order to coincide with the annual morbidity records.

Both datasets used contained figures from 1950 through 1994. Colombia’s mortality statistics were a limiting factor as 1994 was the most recent data available.

The country of Colombia was selected as the study area because of its unique location and subsequent response to ENSO activity and malaria prevalence. Despite the fact that the physical geography is quite varied across the country in terms of altitude and hydrology, Colombia still represented a suitable model for this study, as currently between 18 to 24 million people live beneath 5,500 feet above sea-level, normally a defining limit in altitude for the *anopheles* mosquito (Bouma 1994).

Colombia was also selected due to its economic development status. Researchers have noted that climate variation will influence epidemics mostly in the developing world, as these are the countries most lacking in current vector and disease control (Stone 1995).

A geographic dispersal in disease outbreak would severely affect the peripheral population, who are most lacking in malaria immunity.

Colombia has been recognized as one of the South American countries most lacking in both mosquito abatement and epidemic control.

Figure 1 Colombia's location within the South American continent.
Yet another primary benefit of selecting Colombia as the study area was that, of all South American countries, Colombia has the most complete records of mortality.

Analysis was conducted using linear correlation, a common procedure to measure the relationship between two variables. Linear correlation is an indicator of how closely certain pairs of data fall to a straight line and the correlation coefficient is represented by the letter r. An r-value of +1 indicates that there is a perfect linear relationship between the variables; as one increases, so does the other. An r of −1 indicates that as one variable increases, the other decreases. An r of 0 indicates that there is no statistical relationship between the variables (Utts 1999). Although this type of correlation is only capable of accurately measuring linear relationships, it was expected that among the data used, a generally positive linear relationship would be derived between sea-surface temperature and malaria prevalence.

In this analysis, the two variables were a) annual malaria mortalities in Colombia (January through December), and b) average annual Eastern Equatorial Pacific sea-surface temperature (January through December). Each variable was matched with the other from its respective year.

**Correlation, in terms of r-value, was found with the equation:**

\[
    r = \frac{S_{XY}}{\sqrt{SS_X}\sqrt{SS_Y}}
\]

where \( SS_X = \sum x^2 - \left(\frac{\left(\sum x\right)^2}{n}\right) \)

and \( SS_Y = \sum y^2 - \left(\frac{\left(\sum y\right)^2}{n}\right) \)

and \( S_{XY} = \sum xy - \left(\frac{\left(\sum x\right)\left(\sum y\right)}{n}\right) \)

let \( x = \) yearly average sea-surface temperature
let $y =$ number of cases of malaria reported annually in Colombia

let $n =$ number of pairs of $x$ and $y$
Definition of Terms

The following section is a concise dictionary of terms used within this paper and are defined in particular as they relate to the research conducted. Although these terms are widely common within the fields of epidemiology and climatology, some are more specifically related to malaria and El Niño, respectively.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>anomaly</td>
<td>any specified mathematical deviation from a climate norm</td>
</tr>
<tr>
<td>anopheles</td>
<td>any mosquito of the genus <em>anopheles</em>, which can carry the malaria parasite and transmit the disease to humans</td>
</tr>
<tr>
<td>ENSO</td>
<td>El Niño Southern Oscillation, or El Niño; a warming of the ocean surface off the western coast of South America creating unusual weather patterns in various parts of the world</td>
</tr>
<tr>
<td>geographical climatopathology</td>
<td>the study of the influence of climate on pathology and subsequent geographic distribution of disease</td>
</tr>
<tr>
<td>host</td>
<td>an animal or plant in which a disease lives</td>
</tr>
<tr>
<td>isotherm</td>
<td>line on a weather map connecting all observations of equal or constant temperature within a certain time frame</td>
</tr>
<tr>
<td>malaria</td>
<td>an infectious tropical and subtropical disease characterized by cycles of chills, fever, and sweating, caused by a protozoan of the genus <em>plasmodium</em> in red blood cells, and transmitted to humans by the bite of an infected female <em>anopheles</em> mosquito</td>
</tr>
<tr>
<td>morbidity</td>
<td>the rate of incidence of a disease as measured by the number of reported cases of a disease</td>
</tr>
<tr>
<td>mortality</td>
<td>the rate of deaths caused by a disease for a specified population</td>
</tr>
<tr>
<td>parasite</td>
<td>any organism which lives during the whole or part of its existence on or in the body of another; the causative agent of a disease; in the case of malaria, the <em>plasmodium</em></td>
</tr>
<tr>
<td>pathogen</td>
<td>all agents involved in disease transmission; in the case of malaria, it includes the parasite, vector, and host</td>
</tr>
<tr>
<td>plasmodium</td>
<td>protozoan of the genus <em>plasmodium</em>, the parasite, or causative agent of malaria</td>
</tr>
<tr>
<td>SST</td>
<td>sea-surface temperature, or skin temperature of the ocean surface, as sensed from buoys and satellite remote sensing instrumentation</td>
</tr>
<tr>
<td>three-factor complex</td>
<td>the cyclical transmission of disease involving a parasite, vector, and host</td>
</tr>
<tr>
<td>transmissible disease</td>
<td>any disease capable of being transmitted by infection; synonymous with communicable, contagious, or contractable</td>
</tr>
<tr>
<td>vector</td>
<td>any organism, such as a mosquito, that carries disease-causing microorganism</td>
</tr>
</tbody>
</table>
Discussion of Results

In order to assess the validity of using sea-surface temperatures to forecast malaria outbreaks, plotting the data in a simple linear regression model was necessary. Annual malaria mortalities in Colombia (January through December) were paired with average annual Eastern Equatorial Pacific sea-surface temperature (January through December of the respective year). The raw data was then formulated to find the correlation coefficient.

Figure 2  Example of the dataset constructed for statistical analysis
Indeed, as hypothesized, a positive linear correlation was found between yearly average eastern equatorial Pacific sea-surface temperatures and number of malaria cases reported annually in Colombia. An \( r \)-value of 0.4049 was determined through the aforementioned statistical procedures, indicating that as sea-surface temperature increases, so do the number of reported malaria cases.

Figure 3  A correlation of 0.4049 was found between yearly average sea-surface temperatures and number of malaria cases reported annually.

Despite a few outliers that may have inflated correlation, the data is legitimate. It appears that the explanatory variable of sea-surface temperature is a contributor, but not the sole cause of the response in the number of malaria cases reported. Additionally, confounding variables that cannot be separated from the dataset make valid interpretation a difficult task. In such case, statistical significance is a relative factor dependent upon the validation of each individual researcher. Further studies under various conditions could lead to an increase in reliability of this particular statistical correlation.
Despite the intricacies of statistical validation, the positive relationship seems to confirm that ENSO (indicated by anomalous increase in SST) does promote malaria prevalence through its effects on the parasite, vector, host, or a combination of all three. The most explanatory factor is that El Niño Southern Oscillation affects the vector, the anopheles mosquito. This is because of the anopheles extreme sensitivity to climate in habitat selection, reproduction, and maturation.

In regard to the vector, increased temperatures caused by ENSO could prompt sexual maturation of the anopheles, increasing both reproduction and infectious activity. Furthermore, El Niño’s anomalous fluctuations in precipitation could create a habitat more conducive for mosquito living and breeding. However, further investigations correlating precipitation amount with malaria prevalence would be necessary to make this assumption.

Environmental variables, especially soil moisture, are influenced heavily by ENSO-induced rainfall. In turn, soil moisture and hydrology can affect mosquito habitat. An increase in moisture-laden soil and stagnant waters within a riparian, or riverside environment may be particularly suitable for mosquito habitat (May 1954). More research in this area is needed to form a more accurate conclusion between soil moisture and mosquito habitat.

Of course, the number of environmental variables influencing mosquito habitat, reproduction, and infectious activity are practically limitless. Flooding and drought could promote or discourage mosquito habitat, depending on a number of micro-scale geographic features. However, a general agreement between ENSO-induced temperature and precipitation increase in Colombia suggests that higher sea-surface temperatures are
a conceivable factor behind an increase in mosquito activity, and consequently, malaria prevalence.

Also sensitive to climate is the parasitic plasmodium within the transmitting *anopheles*. It is known that temperature and humidity can affect the extrinsic incubation period of the plasmodium. Plasmodium replication increases drastically between 20°C and 27°C, which would almost always be congruent with ENSO events (Patz 1996). However, the effects of climate are more limiting to the *anopheles* than to the plasmodium itself, and could therefore be excluded as a variable in any further correlation.

Although a 0.4049 does signify a certain degree of correlation, it is not as strong of a correlation as anticipated. This may be the result of a number of factors including the effects of a lag between a rise in sea-surface temperature and associated effects of mosquito habitat, and a lack of temporal resolution in both datasets.

The sequential linkage between sea-surface temperature, ENSO-induced temperature and precipitation, and mosquito habitat and breeding is highly complex. For instance, an increase in SST could promote more convective activity, which would prompt more precipitation, which would cause more flooding. However, there is most certainly a time delay between the first and last of these events. This example of delay is extremely variable and not easily measured, and is not quantifiable within the relatively large unit of a year, the format in which the data is presented. Because of this, there is a high possibility of a lag between any of these multiple stages being reflected in the data. Because of the nature of the statistical analysis conducted, by no means do the results represent any type of continuous temporal data. Further analysis conducted on a more
defined temporal scale could help in understanding if there is indeed a lag between any of these multiples stages of the sequence.

Another alteration that could clarify interpretation of results is the use of datasets with a greater temporal resolution. Because of the temporal availability of data, yearly average sea-surface temperatures had to be computed to match the number of malaria cases reported annually. Anytime data is compromised in this manner, precision is sacrificed and certainly some accuracy is lost. Greater refinement in records of disease morbidity would alleviate this problem.

Also, a greater refinement in terms of spatial data could also improve correlation results. Using smaller, regional or county enumeration units would minimize extraneous variables such as human habitation, mosquito vector ecology, hydrology, coastal environments, and altitudinal zonation. However, micro-scale geographic features, such a subtle topographical changes and soil type, would still confound results slightly.

In conclusion, it appears that sea-surface temperature could be used with moderate reliability to predict the probability of a malaria epidemic. Further extrapolation may even allow researchers to forecast the severity of an outbreak. These methods could then be tested and applied to certain geographic areas around the world.

Overall, this analysis accomplished its general purpose of correlating sea-surface temperature with malaria prevalence in Colombia. However, as always, further research, greater availability and continuity of data, and multiple statistical analyses could improve upon the accuracy of results, thus making an ENSO-based warning system of malaria quite successful.
Conclusion

In attempting to correlate sea-surface temperatures with malaria prevalence, there were some potential pitfalls. Confounding geographic variables such as hydrology, land cover, and local medical availability could not be separated from the spatial study area. Additionally, such factors as lags and poor temporal data resolution limited statistical analysis significantly. Some of these potential problems can be alleviated with further devotion to research.

In order to forecast disease outbreaks it is necessary to identify correlations among the underlying factors of the disease itself. So numerous are these variables that their discovery is based almost solely on trial and error. Thus, epidemic intelligence is a vital component in the science of epidemiology.

One branch of epidemiology, geographical climatopathology, seeks to discern the climatic variables controlling disease pathogens and their geographic distribution. The discipline studies disease on a variety of spatial and temporal scales, usually through statistical applications.

The climatic phenomenon known as El Niño Southern Oscillation is only one of many variables influencing malaria epidemics in Colombia. Even in itself, the climate regime of ENSO is difficult to characterize simply as it varies both spatially and temporally. Hence, there would be potential pitfalls in extrapolating a valid climate-disease correlation from one spatial-temporal scale and applying them to another. Still, understanding patterns of climate and their control over disease dynamics can be extremely helpful in both forecasting epidemics and mapping disease distribution.
In assessing the results of this regression analysis, a stronger linear correlation is necessary to accurately forecast malaria epidemics. Currently, only rudimentary predictions would be formulated. As aforementioned, more exacting research, in combination with more spatially and temporally continuous data, could create a correlation that would feasibly allow precision forecasts to be calculated. Not only would these forecasts be able to predict the temporal aspect of an outbreak, but data could be extrapolated to indicate the prevalence of an epidemic.

With refinement in correlation accuracy, imposing a monitoring method that would utilize sea-surface temperatures to indicate the threat of impending malaria epidemics could be a highly successful warning system. Knowledge of spatial and temporal aspects of an approaching outbreak could be extremely valuable in an effort to couple disease surveillance with prediction and prevention. Ultimately, this would allow the efficient establishment of immunization when and where it is needed most in order to save lives. First, however, global epidemic surveillance must be augmented with greater international communication between national health organizations. Additionally, more manpower devoted to research in these fields would also enhance global epidemic intelligence programs. Such a simple and relatively inexpensive surveillance system of this sort could be designed, governed, and maintained by any one of a number of large international health institutions.

Research concerning the linkages of climate and disease must be strengthened before any operation forecast systems in functional. Further studies could bolster research pertaining to the long-term affects of climate change on vectors and disease distribution. Certain formats of long-term disease mapping and modeling could be
especially effective in understanding climate-disease relationships, as they would allot
time for researchers to develop the medical advancements necessary to prevent an
epidemic before its actual occurrence.
Bibliography


